ON THE APPLICATION OF ARTIFICIAL NEURAL NETWORKS
IN ANALYZING AND CLASSIFYING THE HUMAN CHROMOSOMES

JAFAR S. AL HASANI, IBRAHIEM M. M. EL EMARY,
HANEEN N. HEYASAT AND ALA'A A. AL-ZU'BI

Al- Balqa' Applied University
Prince Abdullah Ben Ghazi College for Science and Information Technology
Al Salt- Jordan
E-mail: *Jafar_alhasani@yahoo.com & omary57@hotmail.com*

ABSTRACT

This paper aimed to represent an advanced software tool assists in analyzing and classifying the human chromosomes. Our new software handles both classifying and analyzing the medical image of human chromosomes according to the following steps: **1-** we use the image processing utilities in Matlab to analyze the image; where we provide a new filter to remove the noise depending on the objects that exists in the image. **2-** The filtered image enters to a segmentation algorithm to segment this image to the most acceptable one. **3-**The segments will enter two tracks for classifying the chromosomes that are contained in the segment: the first one depends on using the image processing for measuring the length of each chromosome; while the second one deals with initiating the neural network to classify the chromosomes according to the chromosomes characteristics. There is a noticeable enhancement in the image processing approach as a result of using our filter which improve the image quality by a factor equal to 91.7%.
Keywords: Matlab, Multi-Layer Feed Forward Neural Network, Backpropagation, Skeletonization, Image Profile and Metaphase.

1. INTRODUCTION
Chromosomes are tightly coiled microscopic rod-like structures of DNA and protein that are found in the nuclei of Eukaryotic cells. Each chromosome consists of two sets of genetic material connected at the centromere. The human body chromosomes are 46 chromosomes, arranged into 23 pairs, 22 for body chromosomes, and one pair for sexual chromosomes [1-3]. This paper got its idea from a problem that faces the technicians in the cytogenetics labs in analyzing the human chromosomes. Such problem can be summarized in: counting the number of chromosomes (taking into consideration overlapping and touching chromosomes), studying the banding patterns in each one, and determining the possible genetic disease, if there is any. So, instead of doing the classification process manually by a microscope with the naked eye, our software can do this task using both Image processing approach and Neural Networks (adapting the Matlab utilities). Our software package can be characterized by minimizing the consuming time needed to classify the chromosomes and gain reliable results.

To accomplish the above objective, we use a digital camera and microscope to capture the metaphase image. Also, we adapt the capabilities of Matlab tool to build our software. As a final conclusion to our works, we implement two tracks to classify human chromosomes: using length of the chromosome and Neural network, each one gives us some types of results.To achieve what has been mentioned as a target, we present this work which is organized into the following sections: section two deals with the related works. In section three, we formulate the problem. Section four describes the proposed software. Matlab as a tool for building our software was explained in section five. Our paper was terminated with section six which was devoted to summery, conclusion and future works.

2. RELATED WORKS
There exists many trials have been done in order to classify the human chromosomes; we will focus on the one that uses Artificial Intelligent especially Neural Network. In these trails, the Neural Network handles the image segmentation, chromosomes classification and rearranges them into pairs.

Seung Yun Ryu, Jong Man Cho and Seung Hyo Woo [4] selected some features of the chromosome like: relative length, normalized density profile (d.p) and centromeric index to be used in identification of human chromosomes. They trained neural network classifier by changing the number of samples which used to get (d.p.) to classify the chromosomes.

Boaz Lerner [5] covered all aspects of human chromosome analysis: segmentation using separation of clusters of partially occluded chromosomes, feature description, selection and extraction and classification by Multi-Layer Perceptron (MLP) Neural Network.

In our paper, we create a connection between image processing and Neural Network, i.e. the image segmentation is done by image processing techniques, and the classification is done by Neural Network.

3. PROBLEM FORMULATION

When technicians in the cytogenetics labs want to examine the human cell to see if there is any genetic disorder, they take a cell sample from the patient and prepare the slide to: study it with a very high magnifying power microscope, count the chromosomes and study the banding pattern in each one of them. In order to get correct results, this process might be repeated many times with different types of cells. Both the time and effort to accomplish these tasks are relatively long which implies that there is a necessary demand for a new approach. So, we present this work to implement the above tasks using an advanced software tool.

4. PROPOSED SOFTWARE APPROACH

The process of human chromosome classification executed by our proposed software passes through two main stages given by the following:-

1- Image processing; this phase is concerned with analyzing the image which is taken from the microscope via digital camera and examine this image if it is in RGB form or not. If it was in that form, we should convert it into gray scale image to make it easier to be handled by Matlab. After that, we apply an advanced filtering technique (since we don't use the built-in filters in Matlab) to remove the noise which is exists in the image. We create our own filter because all types of prefabricated filters are used to remove creation types of noise. But in the treated images (see appendix A), the object which we considered it as noise and must be removed is in fact an object in the image. Several of the built in filters (that exist in Matlab) use the value of an output pixel that is determined by the median of the neighborhood pixels rather than the mean, but in our image we want to separate some pixels from its neighborhood. Our filtering algorithm removed 91.7% of the noise in the image. After that, the segmentation process begins with cutting the image into the most likely segments which usually contain one chromosome. However; some critical segments could contain the overlapping or touching chromosomes.

As a new track in chromosome classification, we calculate the length of each chromosome in a single segment to classify the chromosome according to the length using "Skeletonization / Medial Axis Transformation" technique. According to this calculated length, we can classify the chromosome.

2- Neural Network; it is the second stage of chromosome classification that begins with initiation of the Multi-Layer Neural Network with Back-Propagation algorithm and the suitable architecture form, train and transfer functions.

After creating the Neural Network, the typical chromosomes have been inserted in it to be the target in the training process. The output of the segmentation algorithm, which will be an array of objects, will be converted from matrix form into vector form. This vector represents the new input to the Neural Network.

This Neural Network uses Back-Propagation algorithm. Our work will classify the chromosomes, which means that the neural network will be trained to distinguish them and rearrange them into pairs. Each pair must have two identical chromosomes. If there is any extra or lost or defected chromosome, the neural network will show it through its decision. Also the network will separate the noise which the filter algorithm could not remove and still in the image after segmentation (our proposed flowchart that illustrates these procedures is shown in Fig.1).
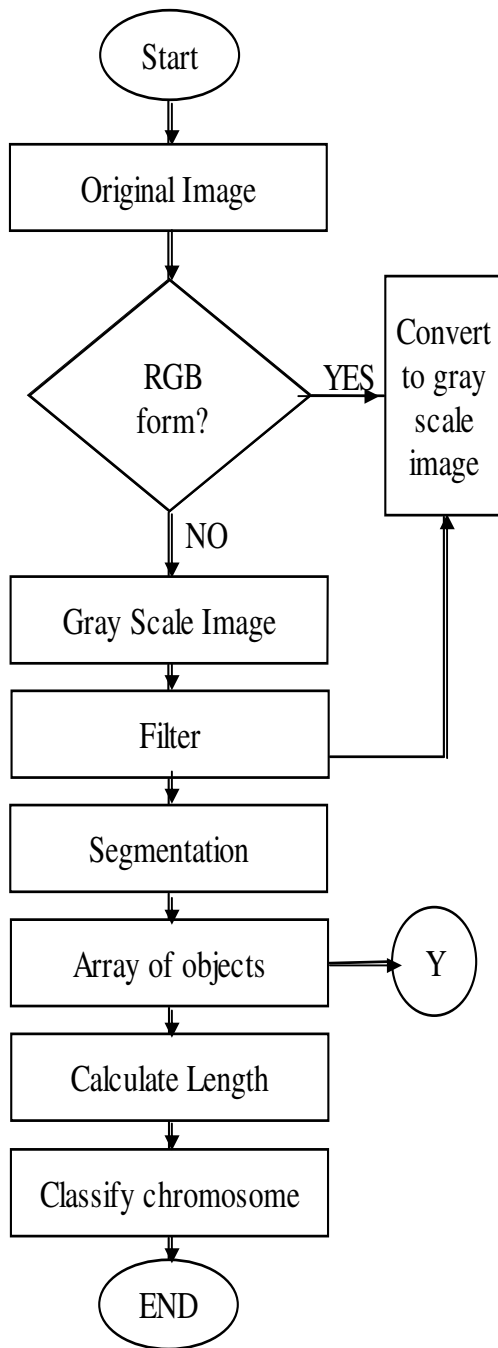
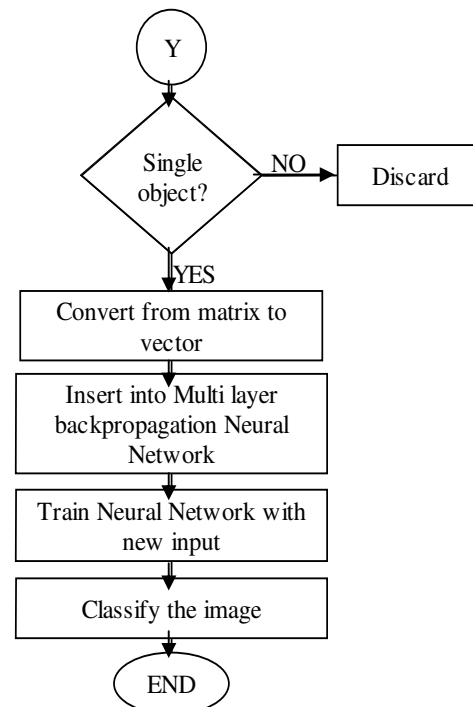**Fig.1 Flowchart of Human Chromosome Classification**

Fig.1 (Cont.)

5- MEDICAL IMAGE PREPROCESSING & ACQUISITIONS

To calculate the chromosome length using Skeletonization / Medial Axis Transformation approach , we apply some Matlab functions to reduce foreground regions in a binary image to skeletal remnant that largely preserves the extent and connectivity of the original region. Each point on the skeleton has an intensity which represents its distance to boundary in the original object. Calculating the length of each object depends on the depth and recursion principle. The calculation proceed as follows: the first phase calculates the number of paths which we have in each object. The second phase takes the tallest path from which we take the length.

Since traditionally the calculated chromosome length was measured by base pair, but we have not this unit in Matlap tool, so we adapt some mathematical transformation given in

[11] to substitute pixel instead of base pair.

6. APPLIED MATLAB TOOLS& OUTPUT RESULTS

In this paper, we use Matlab as a programming language since Matlab is widely used in all areas of applied mathematics in education and research. Matlab stands for MATrix LABoratory and the software is built up around vectors and matrices [6]. This makes the software particularly useful and great tool for solving algebraic and differential equations and for numerical integration. Matlab also has some toolboxes useful for signal processing, image processing, Neural Network,

Database, Wavelet… etc, which have ready built-in functions.

Matlab Image Processing Toolbox contains many powerful tools for image processing. It extends the basic capabilities of Matlab by providing a number of specialized I/O, display and processing functions for image and image processing. When working with images in Matlab, there are many things to keep in mind such as: loading images, using the right format, saving the data as different data types, how to display an image and conversion between different image formats. The following image formats are supported by Matlab: BMP, HDF, JPEG, PCX, TIFF, XWB…etc. The image is stored as matrix using standard Matlab matrix conventions. There are five basic types of image supported by Matlab: Indexed images, Intensity images, Binary images, RGB images and 8-bit images [7- 9].

Artificial Neural Networks (ANNs) as a tool in Matlab are networks of interconnected simple units that are based on a greatly simplified model of the brain. ANNs are useful learning tools by being able to compute results quickly interpolating data well. ANNs are composed of units (also called nodes) that are modeled after neurons, with weighted links interconnecting the units together (see Fig. 2). The main difference between ANNs and other learning mechanisms is that it is composed of these simple units and they work together in a highly parallel manner. The Neural Networks have many algorithms that specify how this network should works.
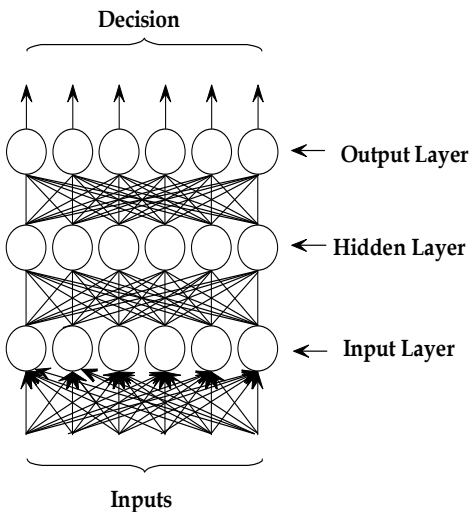
sigmoid function (as shown in Fig.3). This method of training is an example of a supervised learning, where the target of the function is known [10].
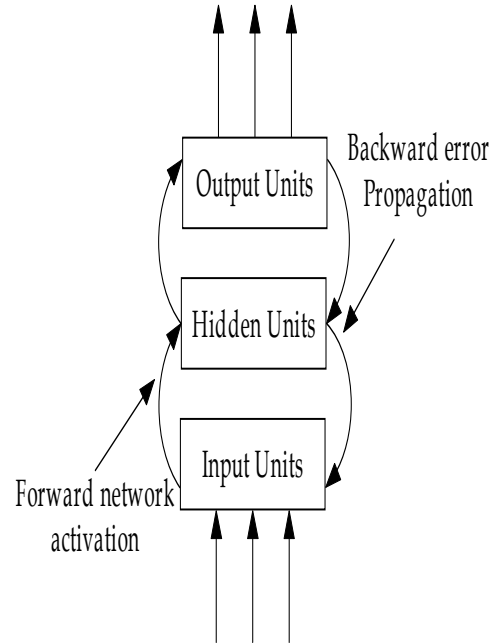

Fig. 3 Back-propagation Algorithm

Fig. 2 Neural Network Architecture

Back-propagation algorithm is the most popular approach to implement learning in neural networks (also called generalized gradient descent and generalized delta rule).This algorithm is a multi-layer network using a weight adjustment based on the

As a result of applying our proposed software in human chromosome classification, we reach to the following results: * The medical image was filtered by 91.7% using our filter.*Classifying the chromosomes can be done through calculating the chromosome length. *Using the Neural Network as another track enables us to distinguish the human chromosomes.

7. SUMMARY, CONCLUSIONS AND FUTURE WORKS

The medical image of human chromosome is treated by our software. It is entered in filter algorithm to remove the noise of the image that it can contain with cleaning percentage reach to 91.7%. Then this filtered image entered to segmentation algorithm, which cuts the original image into most likely segments which can contain one object (one chromosome) or in some bad cases it can contain more than one object (overlapping or touching chromosomes). One of the tracks to classify the human chromosomes is concerned with the length of each chromosome (by converting the image into skeletonization form). Another track of classifying the chromosomes is entering the images after segmentation

into the Neural Network to study each segment and classify the chromosome that is contained in it.

As a future works, we recommend using an expert system to achieve the goals mentioned in this paper and compare its performance with that obtained in our results.

Also, as a tutorial tool, we suggest that the medicine students study this software tool and try to apply it in the classification approach comparing its results with that obtained by their traditional techniques.

REFERENCES

[1]http://www.ornl.gov/sci/techresources/Human_Genome/posters/chromosome/faqs.shtml

[2] Mader, Sylvia S. . BIOLOGY (part 2: Genetic Basics of Life) . 4th edition. Dubuque, Wm. C. Brown Publishers, 1993.

[3] Barch, Margarent J. . The AGT Cytogenetics . 3rd edition . Cambridge, Peter Collin Publishing, 1997

[4] Ryu, Seung Y. & Cho, Jong M. & Woo, Seung H.. A study for the feature selection to identify giemsa-stained human chromosomes Based on Artificial Neural Network
http://bsent.inje.ac.kr/bmetest/bmehtml/boardlink/research/re4_3.pdf

[5] Lerner, Boaz. Toward A Completely Automatic Neural Network Based Human Chromosome Analysis.1998,
http://www.ee.bgu.ac.il/~boaz/IEEESMCLerner1998.pdf

[6] Hanselman, Duane & Littlefield, Bruce. . Mastering Matlab.
Prentice-Hall, 1996.

[7] The Image Processing Toolbox (Version 3) of MatLab 6.5

[8] Gonzalez, Rafael C. . Digital Image Processing. 2[ed] edition. Prentice-Hall, Inc, 2002.

[9] Sandberg, Kristian. Introduction to Image Processing in Matlab.
http://amath.colorado.edu/scico/tutorials/matlab/ImageProcess.html

[10]http://www.cs.sfu.ca/~oschulte/king.pdf

[11]http://www.ornl.gov/sci/techresources/Human_Genome/posters/chromosome/faqs.shtml

Appendix A
Image Before filtration

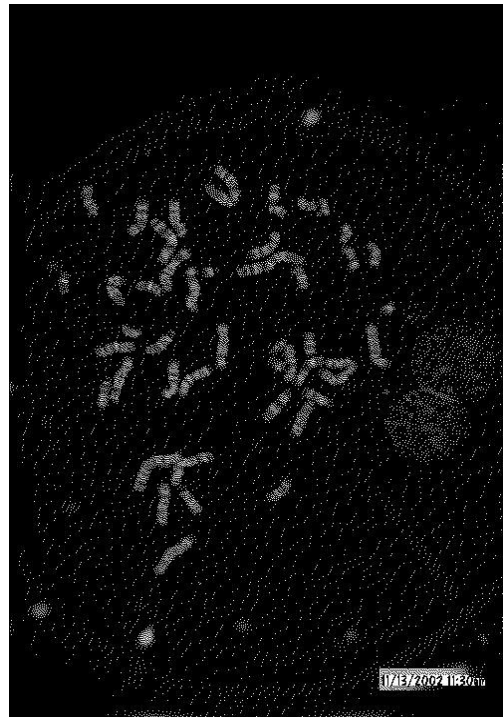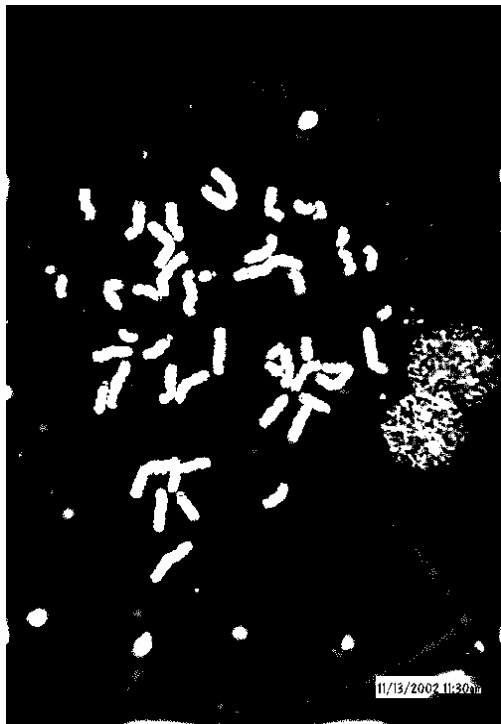*Image After filtration*